

A Bayesian approach to multi-messenger astronomy: Identification of gravitational-wave host galaxies

XiLong Fan^{1,2,3}, Christopher Messenger², and Ik Siong Heng²

1. School of Physics and Electronics Information, Hubei University of Education, 430205 Wuhan, China,

2. SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, United Kingdom

ABSTRACT

We present a general framework for incorporating astrophysical information into Bayesian parameter estimation techniques used by gravitational wave data analysis to facilitate multi-messenger astronomy. Since the progenitors of transient gravitational wave events, such as compact binary coalescences, are likely to be associated with a host galaxy, improvements to the source sky location estimates through the use of host galaxy information are explored. To demonstrate how host galaxy properties can be included, we simulate a population of compact binary coalescences and show that for $\sim 8.5\%$ of simulations with in 200Mpc, the top ten most likely galaxies account for a $\sim 50\%$ of the total probability of hosting a gravitational wave source. The true gravitational wave source host galaxy is in the top ten galaxy candidates $\sim 10\%$ of the time. Furthermore, we show that by including host galaxy information, a better estimate of the inclination angle of a compact binary gravitational wave source can be obtained. We also demonstrate the flexibility of our method by incorporating the use of either B or K band into our analysis.

Subject headings: gravitational waves, parameter estimation, multi-messenger astronomy, electromagnetic follow-ups, sky localisation, Bayesian analysis

1. Introduction

The first detection of gravitational-waves (GWs) will herald the dawn of gravitational wave astronomy and will provide a new way of exploring our universe complementing existing electromagnetic (EM) observations. With gravity coupling very weakly to matter, the detection of gravitational waves is an immense ongoing challenge that pushes both technological and scientific boundaries. Advanced LIGO ([Harry & LIGO Scientific Collaboration 2010](#)) and Advanced Virgo ([The Virgo Collaboration 2009](#)) are expected to have sensitivities that make the detection of GWs a very real prospect in the next few years. Sources of GWs can be classed into 4 broad categories. Continuous GW sources, such as rapidly rotating neutron stars, emit quasi-sinusoidal GWs over durations much longer than the lifetime of the detectors. Stochastic GWs can take the form of a cosmological background, analogous to the EM cosmic microwave background, or could arise from a cacophony

of GW sources at closer distances. Burst GWs are transient signals with poorly modelled or unknown waveforms. Examples of burst sources are supernovae and the merger and post-merger phases of merging compact binaries. Compact binary coalescences (CBCs) are inspiralling binary systems where either or both constituents are a black hole or neutron star. CBCs are the best characterised and one of the most promising sources for the Advanced detectors, with a realistic expected rate of 20 such events per year observed at design sensitivity ([Aasi et al. 2013b](#)).

Multi-messenger astronomy involves the joint observation of astrophysical phenomena using a combination of EM, neutrinos or GW observatories. Examples of multi-messenger astronomy involving GWs include gamma-ray burst observations by Swift ([Evans et al. 2012](#); [Kanner et al. 2012](#); [Aasi et al. 2013c](#)) and Fermi ([Blackburn et al. 2013](#)) and all satellite-based gamma-ray experiments ([Abadie et al. 2012c](#); [Abbott et al. 2010](#)), optical transients by several telescopes ([Aasi et al. 2014](#)) (see a general implementation in [Abadie et al. \(2012a\)](#)) and astroparticle observatories (e.g. high-energy neutri-

³Royal Society Newton Fellow, Xilong.Fan@glasgow.ac.uk

nos, [Bartos et al. 2011](#); [Ando et al. 2013](#)). Such joint observations are likely to be mutually beneficial. For GW observations, an observation in the EM spectrum will allow GW data analysts to focus their searches on a reduced parameter space, thereby improving the sensitivity of their analyses. Conversely, the detection of a GW signal can trigger EM observatories to search for counterpart signals in their respective observation bands. Joint observations will also allow for better characterisation of the signal progenitor, and a richer interpretation of the results of the GW search. For example, searches for GWs in association with GRB051103 ([Abadie et al. 2012b](#)) and GRB070201 ([Abbott et al. 2008](#)) have ruled out the possibility that their progenitors are CBC sources in nearby galaxies. In addition to improving sky location estimates, identifying the host galaxies of GW signal progenitors will enrich this observation by allowing the progenitor environment to be studied which would, for example, provide insight into the evolution of CBC systems. Searches for GW signatures from isolated neutron stars, of both continuous and transient natures, are informed by radio and X-ray observations (see, for example, pulsar glitch and continuous GW searches, [Clark et al. 2007](#); [Hayama et al. 2008](#)). It is also possible to exploit GWs from CBC sources to obtain luminosity distance estimates for source progenitors whilst EM observations of the same event (e.g. gamma-ray bursts) will provide redshift information which can be used to measure the Hubble constant (e.g. [Schutz 1986](#); [Del Pozzo 2012](#)).

Source sky localisation is one of the crucial ingredients for multi-messenger astronomy (e.g. [Sylvestre 2003](#); [Searle et al. 2008](#); [Wen et al. 2008](#); [Wen & Chen 2010](#); [Veitch & Vecchio 2010](#); [Fairhurst 2011](#)). Identifying that a GW signal originates from a sky location consistent with an EM counterpart will establish a clear link between the two observations. Using GW observations to trigger searches for EM counterparts will require accurate and precise estimates of the source sky location to reduce the areas of the sky which the searches for the EM counterpart are to be performed. Uncertainty on the estimated sky location of transient GW signals varies with the strength of the GW signal as well as the location of each GW observatory and their orientations relative to one another. For example, a 3-detector network consisting of the two Advanced LIGO observatories and Advanced Virgo will provide sky localisation estimates of a few tens of square degrees ([Aasi et al. 2013b](#); [Sidery et al.](#)

[2013](#); [Grover et al. 2013](#)) which is a significant challenge for many EM observatories to scan in search for an EM counterpart. In the initial years of Advanced detector operation the sky localisation ability is further impeded by only having the two LIGO detectors (with a slightly less sensitive Virgo after ~ 1 year of operation). The corresponding sky position uncertainties in this case are $O(100 - 1000)$ s of square degrees ([Singer et al. 2014](#)).

An observed GW signal, in particular from CBC sources, provide constraints on both source distance and sky location. With this information, if the host galaxy of the GW signal progenitor can be identified, then EM observations can focus only on the region of sky associated with this galaxy. Amongst the many existing galaxy catalogues, the Gravitational Wave Galaxy Catalogue (GWGC) ([White et al. 2011](#)) has been specifically compiled for current follow-up searches of optical counterparts from GW triggers. [Nuttall & Sutton \(2010\)](#) proposed a ranking statistic to identify the most likely GW host galaxy based on galaxy distance and luminosity and the sky position error box. This ranking method has been adopted in optical follow-up pipeline design ([Nuttall et al. 2013](#)), and optical follow-up observation ([Aasi et al. 2014](#)).

While analyses of Burst GW signals tend not to provide estimates on source distances, due to the assumed unmodelled nature of most burst GW progenitors, it is still desirable to identify potential host galaxies for Burst sources. There are distant-independent algorithms for associating potential host galaxies with observed Burst signals, such as assigning a host galaxy probability based on the surface density of differential number counts of galaxies (e.g. [Bloom et al. 2002](#)).

The expected reach of Advanced LIGO and Advanced Virgo at design sensitivity is ~ 200 Mpc, which is beyond the current range of the GWGC. The detection efficiency improvements for wide-field EM follow-ups obtained through the use of galaxy catalogs has been investigated by ([Hanna et al. 2014](#)). They estimate that an average of ~ 500 galaxies are located in a typical GW sky location error box for NS/NS mergers with Advanced LIGO (~ 20 square degrees), up to range of 200Mpc. By taking into account the GW measurement error in distance and sky location, it was found that the use of a complete galaxy catalogue can improve the probability of successful identification of the host galaxy by $\sim 10 - 300\%$ (depending on the telescope field-of-view) relative to follow-up strategies that do not utilize catalogues.

In the following section we describe the statistical formalism on which our Bayesian approach is based. We define how information from EM observations can be combined with information from GW observations and how this leads to an enhanced ability to identify GW source host galaxies. We also highlight the additional inference power that EM observations lend to the estimation of some GW source parameters, such as the inclination angle of a CBC event. We then describe the specific case of galaxy catalogues representing the EM observation in Sec. 3. In Secs. 4.1 and 4.2 respectively we describe CBC signal waveforms and how, in practice, we combine information from galaxy catalogues into our analysis of GW data and the simulation details. In Sec. 5 we report the results of the CBC simulations that we have performed to validate our method. We conclude with Sec. 6 with a discussion of our results.

2. Inference with joint EM and GW observations

Our aim is to define a method for combining the information contained within EM observations and that obtained through GW observations. In doing so we choose to treat the EM and GW observations and analyses separately up until the point at which parameter estimation on each dataset has been completed. By this we mean that the final output of many EM observations are represented by astronomical results which contain both direct measurements (such as sky position and flux), and estimated values (such as distance and luminosity). The final output of the GW observation is represented by the posterior probability density describing the GW source parameters. Given our model assumption (see Sec. 3) that GW sources are hosted within galaxies there will be common parameters between the two observations, namely the sky position and potentially the distance (dependent upon whether the GW source and/or the EM observation include an estimable distance parameter).

In the following sections we will describe how we combine the information from both observations to enhance our parameter estimation on these common parameters. We also show that due to the correlations between sets of GW parameters, improved knowledge of a parameter that is common between EM and GW observations can also enhance our parameter estimation abilities on non-common parameters.

2.1. Definitions of relevant quantities

To facilitate the formulation of the proposed method for joint EM–GW observations, we start by formally defining the relevant quantities within our problem:

1. The parameter set common to both sets of observations is denoted by γ . In practice in most cases this will consist of the astrophysical source location parameters, the sky position α, β and the distance d .
2. The complete parameter set governing an individual GW event is denoted by $\Theta = (\theta, \gamma)$ which includes the common parameter set γ but also the set θ which does not influence the EM observations. As an example, for a compact binary coalescence source θ could contain (amongst other parameters) the chirp mass, \mathcal{M} . The specific choice of what constitutes a non-common parameter is dependent upon the GW source type and the EM observation.
3. The complete parameter set governing the entire EM dataset is denoted by $\Phi = (\phi, \gamma)$ which includes the common parameter set γ but also the set ϕ which does not influence the GW observations. As an example, if our EM dataset is represented by a galaxy catalogue this could include the galaxy luminosity L_g , metallicity Z , morphology, etc.
4. The GW dataset is denoted by \mathbf{D} potentially consisting of the outputs from multiple GW detectors such that, for the i^{th} detector, $D_i(t) = h_i(t, \Theta) + n_i(t)$, where $h_i(t, \Theta)$ is the GW signal and $n_i(t)$ is the noise from the GW detector. The data can be defined equivalently in the frequency domain.
5. The EM data used is denoted by \mathbf{S} . We do not formally define the specific constituents of \mathbf{S} other than stating that they consist of multiple EM observations.
6. We use M to define our underlying model assumption that links the GW to EM observations. When the EM observation is represented by a galaxy catalogue, then our model assumes GW sources are hosted within galaxies (see Sec. 3).
7. Within our Bayesian framework we use the standard I to contain all additional information.

2.2. Combining GW and EM observations

Our intention is to compute the posterior distribution of the common parameter set γ conditional on both datasets S , D and our underlying model M . Let us start by using Bayes theorem to express the joint distribution on the complete GW parameter set as

$$p(\gamma, \theta|D, S, M, I) = \frac{p(\gamma, \theta|S, M, I)p(D|S, \gamma, \theta, M, I)}{p(D|S, M, I)}. \quad (1)$$

Taking the first term in the numerator we find that

$$\begin{aligned} p(\gamma, \theta|S, M, I) &= p(\gamma|S, \theta, M, I)p(\theta|S, M, I) \\ &= p(\gamma|S, M, I)p(\theta|I) \end{aligned} \quad (2)$$

where we assume that the common parameters are independent of the non-common GW parameters and that our knowledge of the non-common GW parameters are not informed by the EM observations alone.

Here $p(\gamma|S, M, I)$ takes the form of a prior but can also be viewed as the posterior on the common parameters γ as defined by, for example, a catalogue of host galaxies. The quantity $p(D|S, \gamma, \theta, M, I)$ is the likelihood of obtaining the dataset D given S , γ and θ and $p(D|S, M, I)$ is a normalising factor often referred to as the Bayesian evidence.

The likelihood term in Eq. 1 can be simplified by noting that, given γ , the probability of measuring D is fully specified making S redundant. Hence

$$p(D|S, \gamma, \theta, M, I) \equiv p(D|\gamma, \theta, M, I). \quad (3)$$

We can re-express Eq. 3 via Bayes theorem to give us

$$p(D|\gamma, \theta, M, I) = \frac{p(\gamma, \theta|D, M, I)p(D|I)}{p(\gamma, \theta|I)} \quad (4)$$

which we can now substitute back into Eq. 1 to give

$$p(\gamma, \theta|D, S, M, I) = \frac{p(D|I)p(\gamma, \theta|D, M, I)p(\gamma|S, M, I)}{p(D|S, M, I)p(\gamma|I)}. \quad (5)$$

Taking groups of elements in turn we see that there is a constant normalising prefactor equal to a ratio of Bayesian evidences. This is technically a Bayes Factor between two models, the first model being that the GW data contains a GW signal and the second model stating that there is a GW signal in the GW data and it is consistent with the EM observation. It is clearly a function of both GW and EM datasets but is independent of the parameters.

The term $p(\gamma, \theta|D, M, I)$ is the joint posterior probability distribution on all GW parameters obtained from a GW-only analysis. We have, as a denominator, the prior on the common parameters uninfluenced by the EM or GW observations. The reason for its appearance as a denominator is that we must account for the fact that this prior has already been used implicitly twice before, once in constructing the GW posterior and we assume also in the final term, the EM posterior.

To obtain our final goal of a posterior distribution on the common parameters alone we now simply marginalise over the non-common GW parameters in Eq. 1. These parameters are referenced only in the GW-only joint posterior term and hence the posterior distribution on the common parameter set γ conditional on both datasets S and D is

$$p(\gamma|D, S, M, I) = \frac{p(D|I)p(\gamma|D, M, I)p(\gamma|S, M, I)}{p(D|S, M, I)p(\gamma|I)}. \quad (6)$$

Note that again, the common-parameter prior distribution remains as a denominator. We address the practical implementation of this feature in Appendix A.

2.3. Enhanced inference on GW parameters

Additional information from EM observations can be used to improve the inference on non-common GW signal parameters. This ability applies to non-common parameters that exhibit correlation in the GW posterior with one of more parameters in γ . An example of such a parameter is the inclination angle ι in CBC sources which is strongly correlated with the distance parameter d . We stress that this correlation does not exist within the astrophysically motivated prior distribution and is generated by the inclusion of information from the GW dataset.

Using Eq. 5 it can be seen that in general, marginalising over the common parameter set as follows

$$\begin{aligned} p(\theta|D, S, M, I) &= \frac{p(D|I)}{p(D|S, M, I)} \\ &\times \int d\gamma \frac{p(\gamma, \theta|D, M, I)p(\gamma|S, M, I)}{p(\gamma|I)} \end{aligned} \quad (7)$$

does not return a quantity proportional to $p(\theta|D, M, I)$. In this case the EM posterior on the common parameters and the prior in the denominator act to influence the GW posterior to specific regions of the γ space. If the joint GW posterior γ is correlated to any subset of

θ then the inclusion of EM data will have enhanced the inference on these non-common parameters.

In the CBC example, if the common parameters include distance then the EM dataset may identify a possible range or ranges of d that are localised within the GW-only inferred ranges. Having more tightly constrained distance values for a given signal will then correspondingly improve constraints on the inclination angle.

3. Galaxy catalogues as EM observations

Throughout, we work under the assumption that sources of GWs reside within (or in close proximity to) the normal matter seen as galaxies. Sky position and distance estimation from GW observations alone are expected to be relatively uncertain. Hence galaxy catalogues, which we refer to as the EM observation, will help significantly in identifying the GW source host galaxy.

The dependence of the EM data on the common parameter set γ appears clear under the assumption of our general model. If we assume that our galaxy catalogue is complete then there can be no probability of a GW source at any location *not* coincident with a galaxy. We generalise this statement below accounting for the incompleteness of our galaxy catalogues. Additionally, our underlying model may include other parameter dependencies whereby, based on the catalogue alone, one galaxy would be favoured over another. The obvious example parameter in this case is the galaxy luminosity which is strongly related to the galaxy mass and hence the probability of hosting a GW progenitor. Beyond this example we may consider further properties of galaxies that would influence our belief in the presence of a GW source at one location as opposed to another. These may include galaxy type, or metallicity. For future GW searches sensitive to cosmological distances the redshift will influence this belief based on stellar evolution.

All of this information must be encoded into what we call the multi-messenger prior function (MMPF), $p(\gamma|S, M, I)$, where we make clear the inclusion of M representing our underlying model. We interpret the galaxy catalogue location information (sky position and distance) as relating to the probability of the presence of a galaxy, but **not** that it is necessarily related to a GW event. To make this clearer, via Bayes theorem we now decompose this function to give

$$p(\gamma|S, M, I) \propto p(\gamma|S, I)p(M|\gamma, S, I). \quad (8)$$

allowing us to assign this probability based on our understanding of GW progenitor abundance as a function of EM information. The term $p(M|\gamma, S, I)$ represents the probability of our model M (that GW sources are hosted by galaxies) given a specific location and the EM data. The term $p(\gamma|S, I)$ relates to the possible spatial location of the host galaxy and can be inferred directly from the galaxy catalogue. This however requires us to address the issue of the completeness of EM dataset S . Note that the EM dataset that we have (in the form of the galaxy catalogue) if used unmodified will artificially limit the spatial extent of our MMPF. Since the GWGC extends only to 100 Mpc compared with the 200 Mpc sensitive range of the Advanced GW detectors this is certainly the case.

We therefore use an approximation to $p(\gamma|S, I)$ that encodes our ignorance of the galaxy distribution beyond 100 Mpc. In this region $d > 100\text{Mpc}$ we assume a uniform galaxy distribution in volume and hence model the distance prior as $\propto d^2$. We therefore also use an isotropic prior on sky position giving us constant priors on the right ascension and the cosine of the declination. This approximate approach assumes an abrupt transition from the galaxy catalogue prior to spatial ignorance at 100 Mpc and beyond. We therefore implicitly assume 100% completeness of the galaxy catalogue up to 100Mpc. This assumption is incorrect since we know that, the GWGC for example, is only 60% complete beyond 60 Mpc. However, this is a small effect compared to the bias we avoid by taking into account our ignorance beyond 100 Mpc.

To construct $p(\gamma|S, M, I)$, we first consider $p(\gamma|S, I)$ and in doing so we make some assumptions about the sky location of our desired GW signal progenitor with respect to the galaxy properties contained within S . To provide a proof-of-principle of our proposed method, we assume a straightforward form of $p(\gamma|S, I)$ such that

$$p(\gamma|S, I) = \left(\frac{D_{\text{gc}}}{D_{\text{GW}}}\right)^3 \frac{1}{N} \sum_{j=1}^N \delta(\alpha - \alpha_j, \beta - \beta_j, d - d_j) + \frac{3}{4\pi D_{\text{GW}}^3} H(d - D_{\text{gc}}) d^2 \quad (9)$$

where N is the number of galaxies in the catalogue and the subscripted sky position parameters α_j, β_j and d_j denote the galaxy catalogue right ascension, declination and luminosity distance values of the j 'th galaxy respectively. We use D_{gc} and D_{GW} to represent the range of the galaxy catalogue and the sensitive range

of the GW detector network respectively and H is a Heaviside step function. Eq. 9 is constructed to satisfy the constraints that inside the catalogue range only galaxies are considered valid source locations whereas outside, all locations are valid. Also, under the assumption of uniform galaxy distribution in volume the total probability in each region (within the galaxy catalogue range and beyond) is proportional to the volume of that region.

The additional function required to complete the MMPF is the probability of a GW source existing given a location. This is chosen as

$$p(M|\gamma, S, I) \propto \sum_{j=1}^N \delta(\alpha - \alpha_j, \beta - \beta_j, d - d_j) \mathcal{L}_{Bj} + H(d - D_{gc}) \bar{\mathcal{L}}_B \quad (10)$$

where \mathcal{L}_{Bj} is the observed B-band luminosity of the j 'th galaxy reported in the catalogue and $\bar{\mathcal{L}}_B$ is the mean B-band luminosity. As before, the first part of the function describes the known properties of the galaxies within the catalogue and the second part represents our ignorance of the galaxies beyond the catalogue range. In this latter case we assign a GW host probability proportional to the mean/expected B-band luminosity which we obtain from the distribution of luminosities from within the catalogue¹. The final MMPF can then be written as

$$p(\gamma|S, M, I) \propto \left(\frac{D_{gc}}{D_{GW}}\right)^3 \frac{1}{N} \sum_{j=1}^N \delta(\alpha - \alpha_j, \beta - \beta_j, d - d_j) \mathcal{L}_{Bj} + \frac{3\bar{\mathcal{L}}_B}{4\pi D_{GW}^3} H(d - D_{gc}) d^2. \quad (11)$$

This version of MMPF is only a first order approximation to a potentially fully robust way of accounting for catalogue completeness and other EM data effects. Enhancements of the proof-of-principle analysis described in this paper can naturally account for completeness effects due to intrinsic luminosity variation between galaxies, sky position survey sensitivity, etc. It can also allow smooth transitions between our knowledge based on the catalogue and our ignorance beyond the catalogue range. In this case the Heaviside

function in Eqns. 9–11 would be replaced by a more physically motivated function. One can also account for the uncertainties in the measurements of Φ or potential offsets in the observed signal from the centre of its host galaxy (e.g. supernovae kicks) by assigning a distribution to each parameter with a finite variance covering the uncertainties and offsets instead of the Dirac Delta functions used in Eqns. 9–11. While such a function is straightforward to construct mathematically, it will significantly increase the computation cost of our analysis. We leave this for future work but note that our simplistic step-function application of a non-zero MMPF beyond the catalogue range accounts the bulk of the biases imposed by the lack of catalogue completeness.

4. Implementation method

In this section we describe how our approach can be applied to a specific source type in conjunction with our choice of EM observation, a galaxy catalogue. The source we choose is CBC since this allows us to make use the existing lalinference software (Aasi et al. 2013a) for performing the GW inference component of our analysis. We discuss the general application of our method in Sec. 6.

4.1. GWs from compact binary coalescences

The GW signal from a CBC can be divided into 3 parts, the inspiral, merger and ringdown of the final object. For ground-based GW detectors the inspiral stage of binary neutron stars (BNSs) contains the dominant signal-to-noise ratio (SNR) component and the merger and ringdown can be neglected with regards to detection and sky localisation. We can therefore write the inspiral frequency domain waveform using the stationary phase approximation (Dhurandhar & Sathyaprakash 1994) as

$$\tilde{h}(f) = \frac{\mathcal{A}(\varphi, \mathcal{M})}{d} f^{-7/6} e^{-i\{\Psi(f; \mathcal{M}, \eta) - 2\pi f(t_c - \frac{R(\alpha, \beta)^2}{c}) - \phi_c\}}. \quad (12)$$

We define the total mass $M = m_1 + m_2$ and the symmetric mass ratio $\eta = m_1 m_2 / M^2$ where m_1 and m_2 are the component masses. The chirp mass \mathcal{M} is defined as $\mathcal{M} = M\eta^{3/5}$ and d is the luminosity distance of the GW source. The quantity $\mathcal{A}(\varphi)$ is a factor that is determined by the amplitude response of the GW detector and is a function of the chirpmass and the nuisance parameters $\varphi = (\alpha, \beta, \iota, \psi)$ where α and β are the sky position coordinates and ι and

¹For the GWGC catalogue, since we know that it is ~100% complete to within 60 Mpc we take the average luminosity from galaxies within that range.

ψ are the orbital inclination and GW polarization angles respectively. The phase $\Psi(f)$ is a function of the frequency and the mass parameters and in the post-Newtonian point-particle approximation can be Taylor expanded in powers of the dimensionless quantity $(\pi M f)^{2/3}$ (Arun et al. 2005b,a). There is also a frequency dependent phase component with argument proportional to the time of coalescence t_c minus a time delay term representing the GW travel time from a common reference point to the detector location. This time delay is equal to the dot product between the detector position vector \vec{r} and the unit sky position vector \vec{n} . Finally there is a constant phase component ϕ_c , the phase at coalescence.

The common parameters for this galaxy-CBC joint observation are the sky location and distance, giving $\gamma = (\alpha, \beta, d)$. The remaining GW parameters in our analysis $\theta = (m_1, m_2, t, \psi, t_c, \phi_c)$ are marginalised over as described in Eq. 6 and in Appendix A. Reiterating our underlying model, we assume that the GW source resides within a galaxy and hence, following the arguments outlined in Sec. 3 we define the MMPF $p(\gamma|S, M, I)$ using Eq. 11.

4.2. Simulation details

To demonstrate the effectiveness of the proposed method, we simulate a population of GW signals from BNS coalescences. For simplicity, each BNS coalescence is simulated with the exact sky position and distance parameters of a randomly chosen galaxy (we ignore potential supernovae kicks and galaxy catalogue distance uncertainties). The galaxy is chosen from a galaxy catalogue according to an MMPF of the form defined by Eq. 11. We perform 3 separate simulations using a combination of 2 different galaxy catalogues, the initial and advanced LIGO–Virgo GW detector network and different specific choices of luminosity band used in the MMPF. Table 1 gives a summary of our simulations. Note that the MMPF is weighted by B band luminosity for simulations in S1 while K band luminosities are used for S2 and S3 simulations.

The MMPF defined in Eq. 11 allows for the probability of a source being hosted beyond the range of the galaxy catalogue. Whilst we use this MMPF in our analysis of the simulated data, due to computational constraints, we do not simulate *all* required signals beyond the galaxy catalogue range. We perform an equal number of simulations inside and outside the range where we should in fact simulate the fraction

$\approx 1 - (D_{gc}/D_{GW})^3$ beyond D_{gc} . To account for this bias we recycle our costly simulations. This is done by taking the GW posterior samples for each simulation beyond 100Mpc and randomly translating the ensemble of samples in right ascension to a new sky location. For a uniform distribution of sources in volume our MMPF demands that there be on average 7 times as many sources between 100 and 200Mpc as those within 100Mpc. Therefore, each simulation beyond 100Mpc is recycled 6 times. This way we are able to efficiently model the statistical behaviour and variation between these posterior distributions for moderate computational cost.

The first simulations (S1) assumed a 3 network consisting of the two Advanced LIGO detectors and Advanced Virgo. We selected 1000 GW host galaxies from the GWGC (White et al. 2011) according to the first term in the MMPF function (Eq. 11) assuming a complete catalogue up to 100Mpc. And additional 1000 were selected uniformly in volume between 100–200Mpc according to the second term. The latter were then recycled to represent 7000 injections hence closely representing samples drawn directly from the MMPF (Eq. 11). For S1, we follow the approach used in (Nuttall & Sutton 2010) and select host galaxies for simulated BNS coalescence with each galaxy weighted by its B-band luminosity based on Eq. 11. We note that the GWGC is $\sim 100\%$ and $\sim 60\%$ complete out to nearly 40 and 100Mpc respectively and is estimated using the blue band luminosity function ignoring the “Zone of Avoidance”, (see discussion in White et al. 2011). The effects and an appropriate consideration of this type of selection bias will be addressed in future studies.

We also investigate the effects of using a MMPF based on different galaxy properties. It is not clear which choices of astrophysical prior have more effect on GW-galaxy host studies. To that end, the B-band luminosity has been often used as a proxy for the star formation rate of a galaxy. Whether or not the star formation rate is a good tracer of CBC events is much discussed in the literature (e.g. Leibler & Berger 2010; O’Shaughnessy et al. 2010; Fong et al. 2013; Dominik et al. 2013). Among the observable/inferable galaxy properties, the stellar mass of galaxies should be another important (possibly dominating) property to be used. This is because 1) the GW event rate in a galaxy is expected to be proportional to the total number of stars in that galaxy, and 2) the star formation rate is positive

correlated to the stellar mass of star-forming galaxies (e.g. Daddi et al. 2007; Davé 2008; Whitaker et al. 2012; Koyama et al. 2013, and references therein). If the time delay between star-forming and CBC mergers is long enough (e.g. some scenarios in O’Shaughnessy et al. 2010), it is reasonable to believe that the GW CBC-short-duration gamma-ray burst (SGRB) rate in a galaxy is proportional to the old stellar mass of that galaxy, which should be traced by K-band luminosity.

Unfortunately, the GWGC only offers B-band luminosity. For our test of MMPF priors based on different luminosity measures we performed simulations using the Updated Nearby Galaxy Catalog (UNGC) (UNGC, Karachentsev et al. 2013), which contains the most exhaustive list of nearby galaxies properties. This catalogue only extends to ~ 10 Mpc and therefore to avoid issues related to the completeness of the catalogue dominating our results, we use a network consisting of two initial LIGO detectors and Virgo. In this scenario the GW SNRs are comparable to those obtained using sources from the GWGC together with the advanced detector network. To investigate the effects of using a different astrophysical prior function, we have performed BNS coalescence injections by selecting host galaxies from the UNGC based on their K-band luminosity in simulations S2 and S3. We have again made the assumption that the catalogue is complete however, for S2 and S3 we have ignored the volume of space beyond the range of the catalogue. We have therefore assumed an MMPF that contains only the first term in Eq. 11. The primary aim of these simulations is to compare the galaxy ranking effects of using astrophysical prior functions based on the K-band (S2) and B-band (S3) luminosities. Hence our MMPF choice does not influence our results in this case.

In each injection, we generated a GW signal from a BNS coalescences using *lalapps.inspinj*² and injected them into simulated noise from detectors at design sensitivity. Samples from the posterior distribution on the GW parameters are obtained using *lalinference* Aasi et al. (2013a); Veitch & Vecchio (2010) with priors on the signal parameters θ being the same as those used to simulate the signals. In this case the prior distributions on $\cos\iota, \psi, \phi_c$ were uniform across their respective ranges, the priors on m_1, m_2 were assumed uniform on the range $1.3\text{--}1.5M_\odot$ and the time of coalescence was assumed uniform on the time win-

dow $\pm 40\text{msec}$ around the true injected value. The common parameter priors were dealt with according to Appendix A with an assumption that the galaxy catalogue data implicitly incorporates a sky position and distance prior consistent with a uniform distribution in volume.

The final joint posterior distribution is obtained by computing the product of the estimated GW posterior on the common parameters and the galaxy catalogue based EM prior according to Eq. 6. To estimate the GW posterior distribution at the location of a given galaxy using the GW posterior samples, we use a box average method. Namely, we approximate the probability density at a galaxy location as the ratio of the number of samples within a 3D box centred at that location to the total number of samples. We set the sky location size of each box to be $\sim 10\%$ of the range covered by all posterior samples (in both right ascension and declination) and for the box size in distance we choose 10% of the distance of the galaxy in question. This choice of density estimation and the choice of box size does affect the final posterior probabilities at each GW host galaxy candidate but does not strongly affect the relative ranking of density values between galaxies. We plan to address the method of density estimation from posterior GW samples in future work but note that our approach here is adequate in terms of this proof-of-principle analysis.

In general there will be a finite joint posterior contribution from the region of overlap of the GW likelihood and the continuous portion of the MMPF beyond 100Mpc (for S1). The integrated volume of this region represents the joint posterior probability that the source was associated with an unknown galaxy not included in the catalogue. Since such a region is not associated with a specific galaxy we do not include this region in the ranking (i.e. only known galaxies are ranked). However, this probability of the host being outside the catalogue is included when computing the posterior probability of each known galaxy (i.e. if there is a 90% probability of the host being outside the catalogue then the highest possible probability assigned to any known galaxy would be 10%).

²<https://www.lsc-group.phys.uwm.edu/daswg/projects/lalsuite.html>

TABLE 1
SIMULATION AND SEARCH PARAMETERS

Simulation	Injection			GW search parameters
	Galaxy catalogue	Weighting factor	Network	Model
S1	GWGC	B	Adv. L-H-V	B
S2	UNGC	K	initial L-H-V	K
S3	UNGC	K	initial L-H-V	B

NOTE.—For each simulation we indicate the galaxy catalogue (either GWGC or UNGC) and the luminosity-band based weighting factor (B or K) used for simulating sources. The network indicates whether the advanced or initial GW detectors were used and the final column shows the model, based on galaxy luminosity, used by our analysis for each simulation. We note that data generated for the S2 and S3 simulations were identical with only the analysis differing between them.

TABLE 2
THE S1 CASE STUDY SIGNAL INJECTION PARAMETERS

Parameter	Injection value
α	334.7207 deg
β	-1.0587 deg
d	67 Mpc
m_1	$1.40 M_\odot$
m_2	$1.36 M_\odot$
\mathcal{M}_c	$1.20 M_\odot$
η	0.25
$\cos \iota$	0.41
ψ	4.10 rads
Network SNR	46.02

TABLE 3
GALAXY RANKINGS FOR THE SIMULATED SIGNAL PLOTTED IN FIG. 1 FOR DIFFERENT RANKING METHODS

Ranking	Ranking Method		
	Galaxy L_B	Galaxy L_B and d	GW+MMPF (probability)
1	C	C	A (99.36%)
2	A	A	B (0.37%)
3	D	B	C (0.27%)
4	B	E	...
5	E

NOTE.—Blank entries correspond to galaxies with too low probability to rank under each scheme.

5. Results

5.1. Case study examples for S1

5.1.1. Identifying a GW host galaxy

To illustrate our method, we plot a 12 square degree region around the injected sky location of a simulated GW signal in Fig. 1. Table 2 summarizes the injection parameters for this particular signal. The galaxies in this region of the sky are plotted with grayscale asterisk markers with bolder markers representing galaxies with stronger B-band luminosities. Superimposed on the galaxies are contours bounding the 68%, 95%, and 99% confidence regions on the sky location estimate using only the GW data. We note that, by taking the product of the B-band based MMPF (Eq. 11) with the GW sky location posterior following Eq. 6, the top ranked galaxy (A) in this example corresponds to the host galaxy from which the simulated GW signal originates i.e. *the injection host galaxy*. A merit of our approach is that the posterior probabilities assigned to host galaxy candidates can be used to better direct further analyses such as providing a refined sky location for telescopes to observe an EM counterpart signal. In this example, there are three candidates with non-zero posterior probability with the first ranked galaxy having a probability of 0.99 that it is the GW host galaxy.

It is interesting to note that galaxy A in Fig. 1 is only the second most B-band luminous galaxy in this sky region. In fact, the brightest B-band galaxy has been ranked third by our analysis. For a detailed com-

parison, Table 3 breaks down the rankings of the five galaxies in this sky region based on three potential ranking criteria. If we choose to rank galaxies based on B-band luminosity alone, we obtain rankings shown in the second column. Similarly, if we use a statistic that ranks candidate galaxies based on their B-band luminosity combined with the inverse of galaxy distance, a slightly different ranking, shown in column 3, is obtained. In this latter case galaxy D is severely down-weighted due to its distance from the GW sky location estimate. In both of these potential ranking methods the galaxy from which the GW signal originates is not identified as the top ranked galaxy. On the other hand, by combining the GW sky location posterior with a B-band based MMPF (column 4), we identify the host injection galaxy as the top ranked galaxy. Furthermore, by combining the GW sky location posteriors and the B-band based MMPF, two of the five galaxies are not considered likely host galaxies at all. As before, galaxy D is deemed too far away from the bulk of the GW distance posterior and, additionally, galaxy E is too remote from the sky location posterior to be considered significant candidates.

5.1.2. Enhanced inclination angle inference

Once a host galaxy is identified, the distance of the host galaxy can be used to provide an improved estimate on the inclination angle of the CBC progenitor (see Sec. 2.3). With each galaxy in the GW signal error region that is assigned a non-zero probability, a combined posterior distribution on the inclination an-

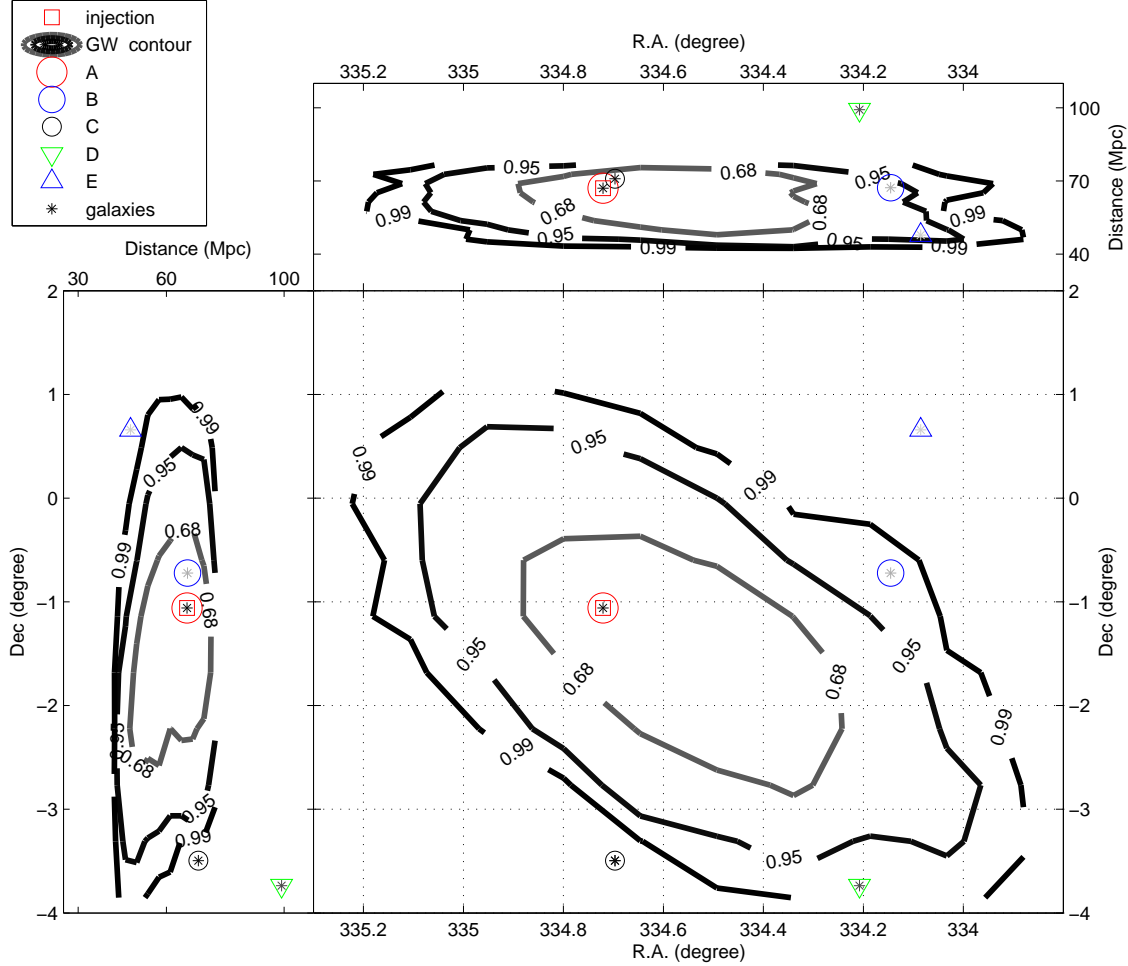


Fig. 1.— Sky localisation for a single BNS coalescence signal with injection parameters shown in Tab. 2 for simulation S1. The contours map out the 68%, 95% and 99% confidence regions on the estimate of the sky location and distance of the signal progenitor obtained using only GW observations. Also plotted are circle markers corresponding to the first (red circle), second (blue circle) and third (black circle) ranked host galaxy candidates, labelled A, B and C respectively, as determined using Eq. 6. Galaxy D (green downward-pointing triangle) has a lower probability because of its distance from the GW sky location estimate. Galaxy E (upward-pointing triangle) is excluded by Eq. 6. Additionally, grayscale asterisk markers are for all galaxies in this sky region, with the shade of the markers corresponding to each galaxies B-band luminosity. The darkest markers are the most luminous in the B-band. The true sky location and distance of the injected BNS coalescence signal (red square) in this case corresponds to the top ranking galaxy candidate. The simulated signal had an optimal SNR of H1: 28.02, L1: 22.67, V1: 28.60, Network: 46.02.

gle can be obtained by multiplying the probably assigned to each galaxy with the posterior distribution on the inclination angle using the GW signal only (see Eq. 7). We plot an example where the inclination angle for the CBC system is improved via our method in Fig. 2. We see that the posterior probability peaks strongly around the correct value when the injection host galaxy is identified as the top ranking candidate. Furthermore, the inclination angle posterior distribution is now much narrower given the additional distance information provided by the galaxy catalogue. In this example, the standard deviation of the $\cos i$ posterior is ± 0.036 after combining the GW posterior with the EM data. Using the GW posterior alone gives a much larger value of ± 0.12 .

5.2. Ensemble statistics for S1

To statistically characterise how well our method, together with the current galaxy catalog adopted in GW astronomy (the GWGC), correctly identifies the host galaxy, we study the injection host galaxy rankings and probabilities for the 8000 simulated GW signals. The 8000 simulations, of which 1000 had hosts selected from the GWGC, produced signals of varying SNR since the source distances ranged up to 200Mpc, the sensitive range of the advanced detector network and source orientation parameters were drawn randomly. Table 4 shows the corresponding SNR distribution for simulation S1.

We choose to focus on the following statistical measures to quantify the effectiveness of our approach. First we define the minimum number of galaxies, N_Z , within the GWGC required to cover a fraction Z of the posterior probability such that $\sum_{j=1}^{N_Z} p_j \geq Z$, where p_j is joint GW-catalogue posterior probability of the j 'th ranked galaxy candidate. In the left panel of Fig. 3 we plot the cumulative fraction of all 8000 S1 simulations required to obtain a desired probability of 0.5, 0.9 and 0.99 summed over galaxies against the number of galaxies in the sum. We find that in more than $\sim 8.5\%$ of all simulations, the top 10 galaxies account for a total probability of 0.5. In fact, the single top ranking galaxy itself has a probability of 0.5 in $\sim 4\%$ of all simulations. Similarly, we see that in $\sim 4.5\%$ and $\sim 3\%$ of simulations, the top 10 galaxies sum up to probabilities of 0.9 and 0.99 respectively.

As an alternative figure of merit, we also consider the number of galaxies within the GWGC, ordered by their ranking, required before the true GW injection

host galaxy is included. As is shown in the right panel of Fig. 3, $\sim 5\%$ of all injected signals are correctly identified as the highest ranking. We must stress that only 12.5% of injections in total had a host galaxy from the GWGC catalogue with the remainder being drawn from a continuous distribution, uniform in volume, beyond 100Mpc. Hence the majority of all injections for which no host galaxy was identified were due to them not being drawn from the GWGC. If we focus on ‘‘loud’’ events, such as events with $\text{SNR} > 30$, we find that there is a $\sim 70\%$ probability that we find the true GW host galaxy within the top 3 ranked galaxies. For quieter events, such as those with $10 < \text{SNR} < 30$, there is $\sim 8\%$ chance that the top 10 galaxies would include the GW host galaxy. For very low SNR (< 10) events, there is almost no chance ($\sim 0.1\%$), that the highest ranked galaxy is the true host of the GW source.

To investigate the efficiency of our method at identifying host galaxies, we examine the behavior of simulations that originate from a galaxy in GWGC. In Fig. 4, the central scatter plot shows the SNR versus the posterior probability obtained for the true host galaxy of the injected signal in . We stress that each point on the plot corresponds to these values for each of the 1000 GWGC based S1 simulations and we do not include events injected from beyond the GWGC. To the top and left of the central plot are multiple histograms showing the distribution of the SNRs and injection host galaxy probability for various injection host galaxy ranks.

For optimal network $\text{SNR} > 20$ injections, the top ranking of the injection host galaxy has only a weak dependence on the SNR (see the ranking depended SNR histogram in Fig. 4). However, some injection host galaxies for high SNR events are not ranked first due to the presence of nearby galaxies with greater B-band luminosities or the GW sky location posterior being peaked away from (but still consistent with) the injection host galaxy. Furthermore, we note that approximately 10% of the injections were not assigned a posterior probability since these sources yielded no GW posterior samples within our galaxy-centred sample boxes used for density estimation. We classify these injections as being particularly poorly localised via the GW observation and have been identified as due primarily to low SNR injections as indicated by the yellow (cross) markers in the scatter plot in Fig. 4.

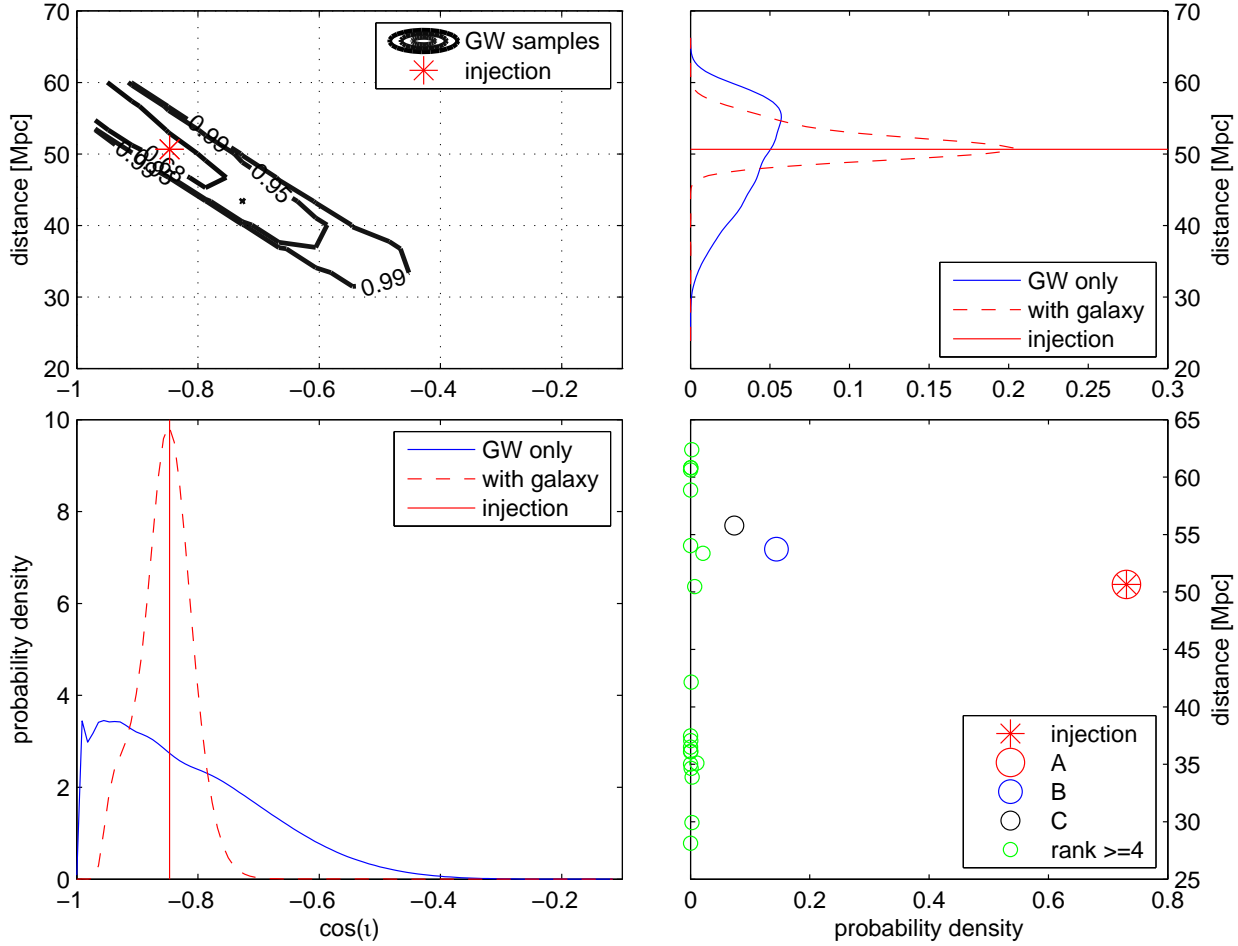


Fig. 2.— An example showing the reduction in the degeneracy between distance and inclination angle ι . The top-left panel shows GW posterior samples. The lower-right panel corresponds to the marginalised posterior probability of each host galaxy as a function of its distance. The circle markers correspond to the first (red), second (blue) and third (black) ranked host galaxy candidates, labelled A, B and C respectively, as determined using Eq. 6. The top-right and lower-left panels correspond to the marginalised probability density functions on distance and $\cos \iota$, respectively. The small-sharp features seen in the solid blue GW-only curve in the lower left-panel are artifacts of the smoothing function used to convert samples to densities. Solid blue and red dashed curves correspond to using only GW posterior samples and GW-galaxy catalogue information, respectively. The simulation injection values of distance and $\cos \iota$ are shown as solid red lines. The simulated signal had an optimal SNR of H1: 44.70, L1: 39.32, V1: 27.26, Network: 65.48.

TABLE 4
FRACTION IN OPTIMAL NETWORK SNR BINS FOR EACH SIMULATION

SNR bins	fraction		
	0-10	10-30	>30
S1	0.295	0.637	0.068
S2 (S3)	0.149	0.586	0.265

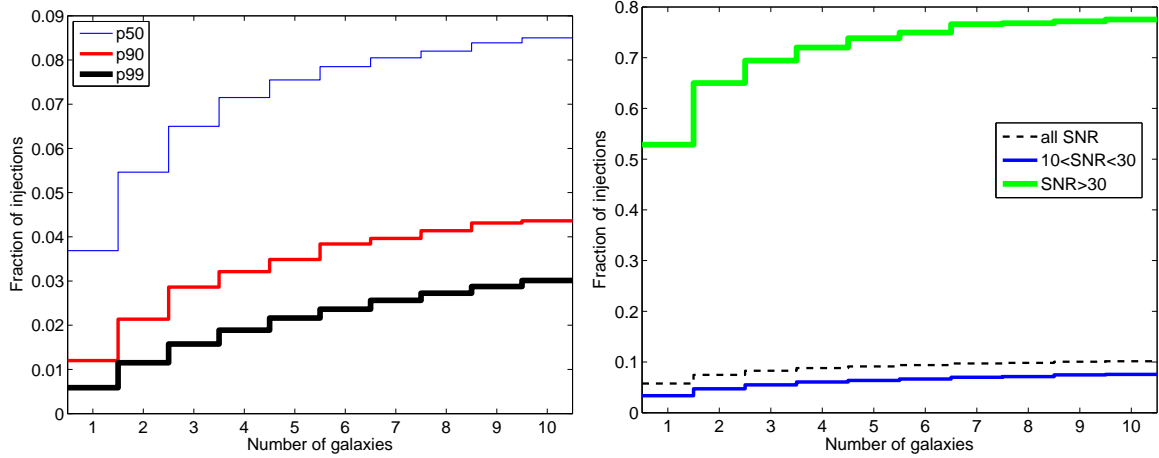


Fig. 3.— Left panel: the fraction of analysed signal injections to have 50%, 90% and 99% probability versus the minimum number of galaxies within the GWGC required to achieve that probability for the S1 simulation. For example, $\sim 4\%$ of the time, the top 7 ranked galaxies would contain 90% of the posterior probability. Right panel: the fraction of analysed signal injections to contain the true host galaxy at or above a particular ranking as a function of that ranking for the S1 simulation. Different curves are shown for various SNR ranges. For example, 10% of the time, the true host galaxy is ranked 10th or higher when considering signals of all SNRs.

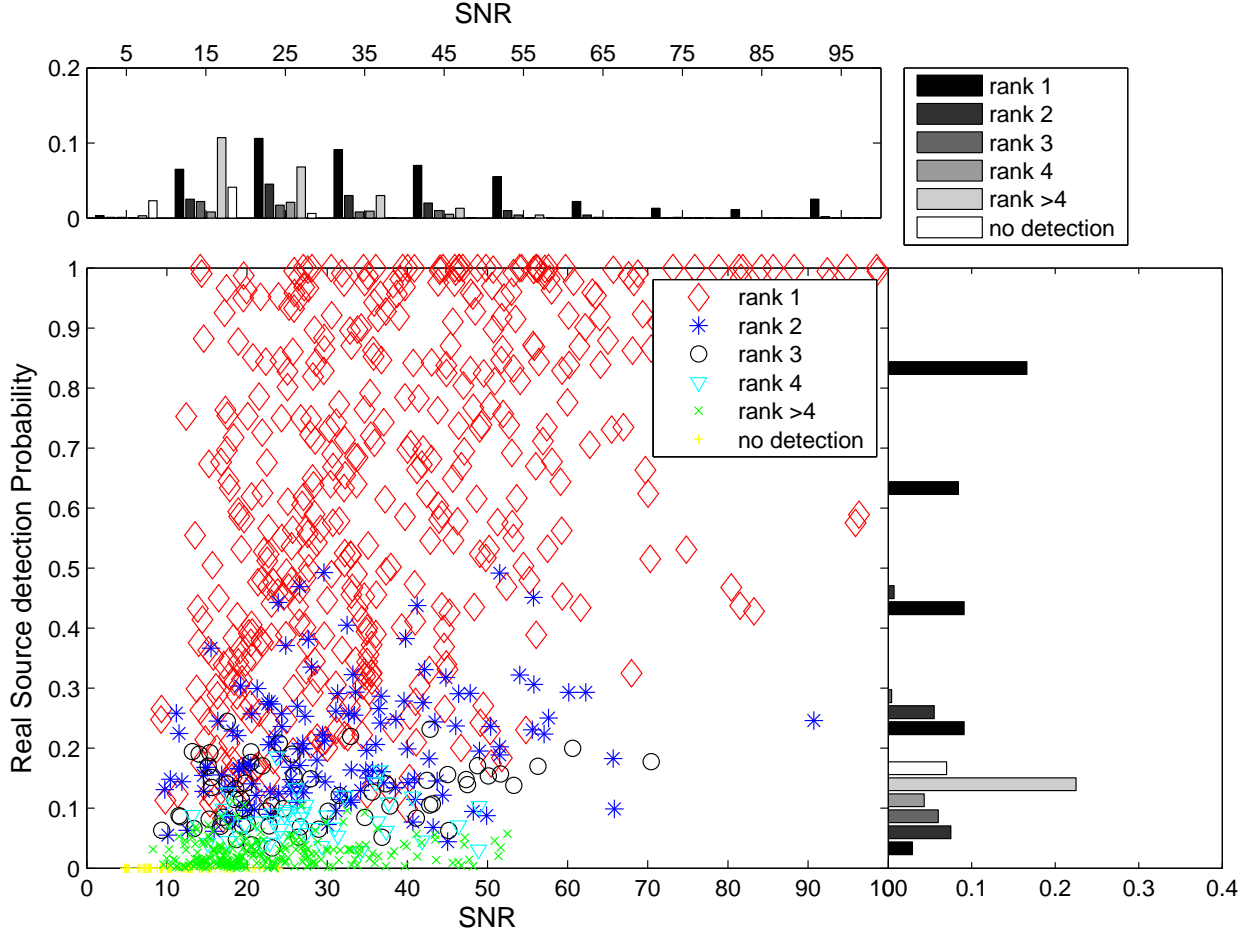


Fig. 4.— A scatter plot showing the probability of the combined GW-galaxy catalogue signal posterior versus the injected SNR for the 1000 simulation in S1 that originated from a GWGC galaxy. Plotted above the central scatter plot are the fractions of all 1000 injections per rank for various SNR ranges. A similar histogram is plotted to the right of the central plot showing the rank distribution for various ranges in injection host galaxy probability. We note that most simulations originating from galaxies in the GWGC catalogue result in signals with an optimal network SNR greater than 10. For these simulations, all galaxies with a probability of 0.2 or more are ranked top 3 or better. Some injection host galaxies for high SNR signals are not identified as the top ranking galaxy due to a combination of nearby galaxies with greater B-band luminosities and GW sky location posteriors being peaked away from the injection host galaxy. About 10% of all simulations are classified *no detection* because a posterior probability was not assigned to the injection host galaxy since no GW posterior samples were found within our galaxy-centred sample boxes. Almost all such simulations had signals with SNR less than 20.

5.3. Using a different MMPF

In the simulation described above the sky location of the injected CBC signals were injected *and* recovered using our approach with a B-band luminosity based MMPF. In simulations S2 and S3 injections were performed using K-band luminosities to determine the relative galaxy MMPF. In simulation S2 the same K-band based MMPF was used for recovery and in simulation S3 the effects of using the B-band for recovery have been studied. In both simulations the actual injection host galaxies were identical, i.e. the same realisation. These 2 simulations have been performed using the initial LIGO-Virgo detector network at design sensitivity since, of our chosen catalogues, the UNGC contains both B and K-band information but has a range of only ~ 10 Mpc.

The colour index, defined by $< B - K >$, depends on galaxy morphology (e.g. Jarrett et al. 2003). Therefore, for K-band luminosity based injections, B-band luminosity together with galaxy morphology could be used as a proxy for the K-band to recover the injection within our MMPF. However, the details of such an analysis are the subject of ongoing study within the multi-messenger astronomy community. Instead, we use simulation S3 to show the effect of using B-band luminosity information for signals simulated using K-band information within our MMPF.

5.3.1. Case study example

Using the “correct” MMPF should lead to a better rank for the injection host galaxy when considering a population of sources. In this section we give an example where using the correct MMPF, by which we mean the same MMPF as that used for simulating the GW signal, leads to a better ranking for the injection host galaxy (see Fig. 5). In particular, the example K-band injection we have chosen ranks as the first in S2, our K-band based recovery MMPF (shown in left panel of Fig. 5) and ranks second when using the B-band based recovery MMPF (shown in right panel Fig. 5). We also note that the 3 highest ranking galaxies in both simulations only share 2 common galaxies. The third rank galaxy estimated using the K-band based recovery MMPF lies within the distance marginalised GW posterior $1-\sigma$ sky map contour, while the third ranking galaxy using the B-band based recovery MMPF lies outside the GW posterior $1-\sigma$ sky map contour. The sky separation of the first and the third candidates returned by the B-band based recovery MMPF (the

“non-correct” one) may challenge an EM follow-up team in terms of observing the top 3 candidates given the limited field of view of EM telescopes.

5.3.2. Ensemble statistics

Distributions of the ranking of the injection host galaxy are shown in Fig. 6. There are 22 of 1000 injections that have a different rank depending on whether the recovery MMPF is a function of the K or B-band luminosity. The K-band based recovery MMPF adopted in S2 and B-band based recovery MMPF adopted in S3 return similar statistical results on injection rank (Fig. 6). Therefore, for this particular galaxy catalog, one can safely use B-band to construct the MMPF when K-band data is not available, although, as one would expect, for some particular cases the K and B-band based MMPF show different results (e.g. see Fig. 5).

6. Conclusions and discussion

We have constructed a Bayesian approach to multi-messenger astronomy and described a proof-of-principle analysis for this approach. The analysis chosen in this case was designed for joint EM and GW observations, in particular, galaxy catalogues and GW events from CBCs. The aim of this research is to improve the sky localisation of GW events by identifying the GW source host galaxy. Identifying the GW host galaxy is, for example, a vital component in using GW signals as cosmological standard-sirens (e.g. Schutz 1986; Del Pozzo 2012) as well as a key ingredient for follow-up observations for EM counterparts to GW signals (e.g. Kanner et al. 2008; Veitch & Vecchio 2010; Nissanke et al. 2013). The proof-of-principle analysis presented here has been demonstrated using simulated BNS events, with their sky locations and distances randomly selected from the GWGC and uniform distribution for number density (in S1) and UNGC (in S2 and S3). Additional prior galaxy weighting was based on individual galaxy measured/inferred parameters and simulated GW signal injections were added to LIGO-Virgo simulated noise. To ensure the SNR distributions are sensible and comparable between choices of catalogue, we have adopted advanced and initial LIGO-Virgo design noise performance when using the GWGC and UNGC related simulations, respectively.

In addition to the standard galaxy candidate ranking, one merit of our approach is that the posterior

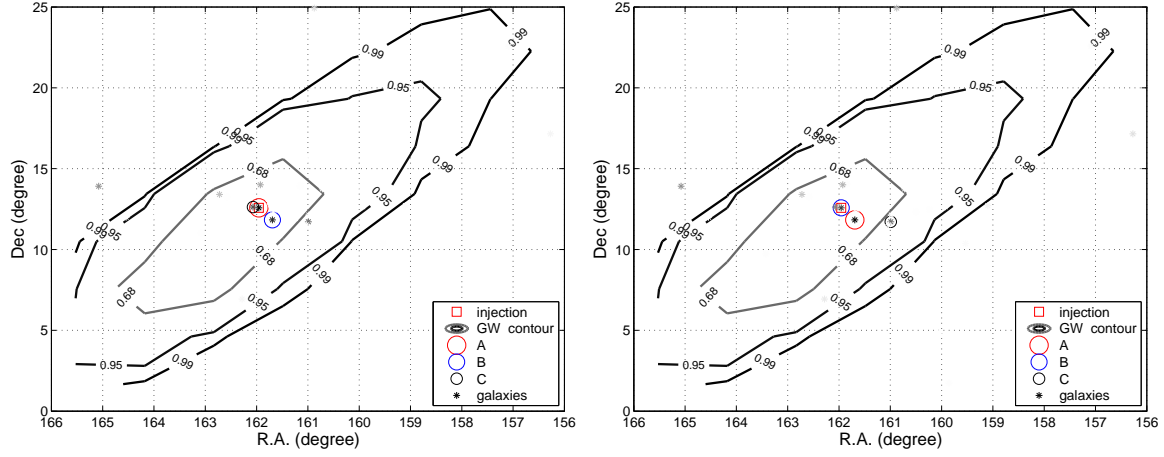


Fig. 5.— Comparison of the first three GW host galaxies in rank recovered using in a K-band based MMPF in simulation S2 (left panel), and in a B-band based MMPF in simulation S3 (right panel). Symbols have the same meaning as defined in Fig. 1. The simulated signal has an SNR of H1: 8.19, L1: 9.92, V1: 6.91, Network: 14.60.

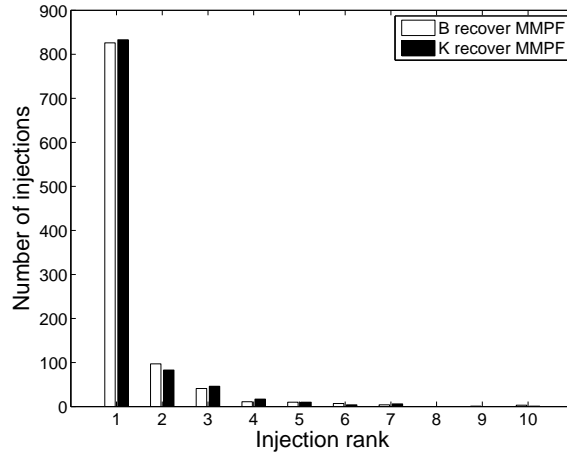


Fig. 6.— Comparison of GW host galaxy rank in S2 and S3 simulations. The K-band based recover MMPF adopted in S2 and B-band based recover MMPF adopted in S3 return similar statistical results on injection rank.

probability of hosting the GW source associated with each galaxy is also produced. This information could, for example, be used to better guide EM follow-up observations to focus on the particular galaxies with high posterior probability (see case study in Sec. 5.1.1) or to simply limit the galaxies followed-up to those accounting for the bulk of the total probability (see case study in Sec. 5.3.1). Moreover, we have shown that better constraints on the common parameters of the EM–GW observations can enhance the inference on non-common GW parameters. A case study on the improvement on the inference of the inclination angle of CBC events was present in Sec. 5.1.2.

Using 8000 S1 simulations, we found that about 8%, 4% and 3% of injections have 50%, 90% and 99% of the probability included in the top 10 ranked galaxies in GWGC, respectively. The first ranking galaxy has a 50% probability of being the true GW host galaxy in about 4% of injections. These results are dominated by the GWGC 100 Mpc distance cut comparing with the expected reach of Advanced LIGO and Advanced Virgo at design sensitivity ~ 200 Mpc. A deeper and all-sky galaxy catalog is necessary to improve the identify of GW host galaxy.

Although for some particular cases the “correct” K-band based recovery MMPF shows better results (e.g. see case study in Sec. 5.3.1), the K and B-band based recovery MMPF return statistically similar results on injection rank for 1000 injections with a K-band based MMPF using the UNGC. This may result from the fact that 1) the red-sequence galaxies represent only 17% of the total number of galaxies in the local Universe and, 2) it is rare to have a massive red-sequence (no or little ongoing star-formation) galaxy in close proximity to a massive blue cloud (star-forming) galaxy.

We note that the proposed method for incorporating astrophysical information is very flexible. It is straightforward to incorporate a different or updated galaxy catalogue into the MMPF by updating the sky location and distance parameters in Eq. 9. A new, more sophisticated MMPF can also be taken into account by updating Eq. 10. In both cases, there is no need to perform the potentially time consuming step of re-analysing the GW data since the MMPF is constructed to be independent of GW observations. It is just a matter of constructing the MMPF using Eq. 11 and multiplying the new MMPF with the GW posteriors.

Faint galaxies are a challenge for any astronomical survey and therefore also an obstacle in identifying a GW host galaxy using any available galaxy cat-

alogue. The completeness of a catalog is a complex concept and not particularly well defined. Analysis of luminosity functions can give an indication of the level of incompleteness in a catalogue. In White et al. (2011) they estimate the completeness of the GWGC as a function of distance by comparing blue band data with an analytical Schechter galaxy luminosity function within 100 Mpc. In this paper, we only take into account the effect of galaxy catalog distance cut on identifying GW host galaxy in Eq. 9, which is the major task to identify GW host galaxy by GWGC in advanced detectors era. Further effects by the EM data, such as the completeness of the catalog within distance cut and the uncertainty of EM measurements, will be studied in our future work.

One of the major components of our approach is the MMPF, within which we distill our astrophysical knowledge related to our underlying model in Eq. 10, that our GW source resides within a galaxy. In past searches, the B-band luminosity of galaxies is adopted within the GW community to estimate the CBC event rate in galaxies for proposed galaxy catalog based EM follow-up observations (e.g. Nuttall & Sutton 2010; Evans et al. 2012; Aasi et al. 2014). The nature of compact binaries suggests that GWs from CBCs would most likely be hosted by the older stellar population in more massive galaxies. Stellar masses are mainly determined by observed stellar light through the stellar mass-to-light ratio or fitting the spectral energy distribution of galaxies (e.g. Fan et al. 2013), which varies according to a few parameters (see a recent review Courteau et al. 2014, and references therein). Therefore, multi-band luminosity could benefit GW-galaxy host research. Debate pertaining to the reliability of optical and near-infrared stellar mass estimates is currently ongoing. However, the fact that the stellar mass-to-light ratio varies less in near-infrared bands than in blue bands over a wide range of star-formation history (e.g. Bell & de Jong 2001), and old stellar populations (≥ 2 Gyr) are mostly bright in the near-infrared band (e.g. Maraston 1998), suggests that the near-infrared band (e.g. the K-band) is a better tracer of old stellar-population mass, and therefore the CBC derived GW event rate. It is believed that the morphology and metallicity of a galaxy will affect its CBC event rate (e.g. Belczynski et al. 2010; Fryer et al. 2012; O’Shaughnessy et al. 2010) and its long-duration gamma-ray burst (LGRB) (possibly associated with GW bursts) event rate (e.g. Fan et al. 2010). We note that our observations from using

UNGC are dominated by the very low galaxy density at 10 Mpc. Therefore, the selection effects arising from this low density must be taken into account which, as previously mentioned, is the scope for future work. Nonetheless, just as we have started to do with B and K-band luminosities, it is important for future multi-messenger analyses to investigate the influence of this information and/or lack thereof in their simulations.

The density of galaxies around the injection host galaxy and the size of the GW posterior on the signal sky location have a significant impact on the galaxy rankings we have observed. Therefore the galaxy environment may also play a role on identifying GW host galaxies. Galaxy clusters usually have a massive galaxy surrounding less massive galaxies. It will be interesting in the future to test the efficiency of GW host galaxy identification for different GW sources such as LGRBs associated GW burst sources, which are believed to be preferentially hosted by faint (irregular) galaxies (e.g. [Fynbo et al. 2012](#); [Fan et al. 2010](#)).

Beyond our underlying model, that the GW source resides within a galaxy, the potential offset in the observed a GW signal from the centre of its host galaxy (e.g. by supernovae kicks) has been suggested by population models (e.g. [Bloom et al. 1999](#)) and observed SGRB offsets (e.g. [Fong & Berger 2013](#)). This effect could be accounted for using a model that assigns a distribution to common parameters (sky location and distance) covering the offsets. The potential offsets, which are \sim a few to tens of Kpc, should have minimal effect on the GW host galaxy identification even with the lack of faint galaxies in the catalog. This is attributable to the fact that the offset SGRBs are unlikely to be hosted by the unobserved faint galaxies which are far from the SGRB hosts (e.g. [Boylan et al. 2014](#)).

Amongst other dependencies, the potential implications of our Bayesian approach are a function of the common parameters between the two different observation sets. There is much discussion regarding the possible non-GW signatures of BNS mergers (e.g. [Metzger & Berger 2012](#); [Gao et al. 2013](#), among others) and BNSs are commonly accepted as the central engine for SGRBs (e.g. [Paczynski 1986](#); [Eichler et al. 1989](#); [Fox et al. 2005](#)). The opportunity to perform multi-messenger astronomy by observing the SGRB counterpart to the GWs emitted is one of the many reasons that CBC systems are considered an interesting source. Furthermore, weaker kilonova optical counterparts are also expected to be emitted by CBCs (e.g.

[Metzger et al. 2010](#)). While SGRBs are expected to be highly beamed, the kilonova signal radiates isotropically. Direct detection of GWs in coincidence with their SGRB or kilonova counterparts will provide the strongest evidence that SGRB progenitors are merging CBC systems. Besides the sky location and distance (if available), the common parameters between EM and GW signals in these cases could also include the arrive time and source energy which would then be incorporated consistently into the MMPF design.

Precise sky localisation and the consequent host galaxy identification is of prime importance to the most promising and well established ideas in GW cosmology. In ([Schutz 1986](#)) and ([Del Pozzo 2012](#)) the idea was proposed and investigated that correctly combining the potential host galaxy redshifts with the luminosity distance inferred from GW observations would allow measurement of the Hubble constant using first and second generation GW detectors. Improved host galaxy identification such as the method we propose would directly impact and reduce the statistical noise inherent to this cosmological measurement.

As a final remark we consider the imminent GW detection era and the potential 10s–100s of CBC signals detectable with the advanced network of detectors. In such a scenario our approach could easily be inverted to ask a different question, what is the true MMPF? With multiple GW detections it would then be possible to perform model selection on different choices of MMPF allowing the GW data to feedback population information to the wider astrophysical community.

We would like to acknowledge valuable input from J. Kanner, M. Hendry, P. Raffai, and our anonymous referee whose input has greatly improved the manuscript. The authors also gratefully acknowledge the support of this research by the Royal Society, the Scottish Funding Council, the Scottish Universities Physics Alliance and the Science and Technology Facilities Council of the United Kingdom. XF acknowledges financial support from National Natural Science Foundation of China (grant No. 11303009). XF is a Newton Fellow supported by the Royal Society and CM is a Lord Kelvin Adam Smith supported by the University of Glasgow.

A. Generating GW posterior samples

In this section we describe the practical procedure used to generate samples from the GW posterior $p(\boldsymbol{\gamma}|\mathbf{D}, M, I)$. In order to avoid the step of dividing this distribution by the prior on the common parameters (see Eq. 6) we instead effectively sample from the ratio of posterior and the prior. Using Bayes theorem and the assumption that the joint prior on the common and non-common GW parameters is separable such that $p(\boldsymbol{\gamma}, \boldsymbol{\theta}|I) = p(\boldsymbol{\gamma}|I)p(\boldsymbol{\theta}|I)$ we can express this ratio as

$$\frac{p(\boldsymbol{\gamma}, \boldsymbol{\theta}|\mathbf{D}, M, I)}{p(\boldsymbol{\gamma}|I)} = \frac{p(\mathbf{D}|\boldsymbol{\gamma}, \boldsymbol{\theta}, M, I)p(\boldsymbol{\theta}|I)}{p(\mathbf{D}|I)} \quad (\text{A1})$$

This assumption is valid for our GW-galaxy catalogue simulations where the sky position and distance priors are independent of the mass and orientation parameters of the CBC.

Our expression now has explicit priors on $\boldsymbol{\theta}$ but in practice, in order to generate a posterior distribution using existing algorithms, we are required to specify a prior on all parameters within the problem, including $\boldsymbol{\gamma}$. In the GW analysis case this means specifying uniform dummy priors on $\boldsymbol{\gamma}$ such that the function from which samples are drawn is actually

$$X(\boldsymbol{\gamma}, \boldsymbol{\theta}) = \frac{1}{\mathcal{V}_{\boldsymbol{\gamma}}} p(\mathbf{D}|\boldsymbol{\gamma}, \boldsymbol{\theta}, M, I)p(\boldsymbol{\theta}|I) \quad (\text{A2})$$

where $\mathcal{V}_{\boldsymbol{\gamma}}$ is the volume of the common parameter space and its inverse is the uniform prior on $\boldsymbol{\gamma}$. Therefore we specify non-standard priors for the sky position and distance within our sampling algorithm since we assume that the EM observation already contains these priors.

Substituting this into Eq. 5 gives us

$$p(\boldsymbol{\gamma}, \boldsymbol{\theta}|\mathbf{D}, \mathbf{S}, M, I) = \frac{p(\mathbf{D}|I)}{p(\mathbf{D}|\mathbf{S}, M, I)} p(\boldsymbol{\gamma}|\mathbf{S}, M, I) \mathcal{V}_{\boldsymbol{\gamma}} X(\boldsymbol{\gamma}, \boldsymbol{\theta}). \quad (\text{A3})$$

From this point we can proceed as described in Secs. 2.2 and 2.3 whereby terms in the joint EM-GW posterior are marginalised over a subset or all of the non-common parameters. We have therefore made sure that the correct physical priors on the non-common parameters have been applied and we have not over-applied the common parameter priors.

REFERENCES

- Aasi, J., et al. 2013a, *Phys. Rev. D*, 88, 062001
- . 2013b, *ArXiv e-prints*, 1304.0670
- . 2013c, *Phys. Rev. D*, 88, 122004
- . 2014, *ApJS*, 211, 7
- Abadie, J., et al. 2012a, *A&A*, 539, A124
- . 2012b, *ApJ*, 755, 2
- . 2012c, *ApJ*, 760, 12
- Abbott, B., et al. 2008, *ApJ*, 681, 1419
- Abbott, B. P., et al. 2010, *ApJ*, 715, 1438
- Ando, S., et al. 2013, *Rev. Mod. Phys.*, 85, 1401
- Arun, K. G., Iyer, B. R., Sathyaprakash, B. S., & Sundararajan, P. A. 2005a, *Physical Review D*, 72, 69903
- . 2005b, *Physical Review D*, 71, 84008
- Bartos, I., Finley, C., Corsi, A., & Márka, S. 2011, *Physical Review Letters*, 107, 251101
- Belczynski, K., Dominik, M., Bulik, T., O’Shaughnessy, R., Fryer, C., & Holz, D. E. 2010, *ApJ*, 715, L138
- Bell, E. F., & de Jong, R. S. 2001, *ApJ*, 550, 212
- Blackburn, L., Briggs, M. S., Camp, J., Christensen, N., Connaughton, V., Jenke, P., & Veitch, J. 2013, *ArXiv e-prints*:1303.2174, 1303.2174
- Bloom, J. S., Kulkarni, S. R., & Djorgovski, S. G. 2002, *AJ*, 123, 1111
- Bloom, J. S., Sigurdsson, S., & Pols, O. R. 1999, *MNRAS*, 305, 763
- Boylan, C., Li, Y., Fan, X. L., & Heng, I. S. 2014, *ArXiv e-prints*, 1401.7851
- Clark, J., Heng, I. S., Pitkin, M., & Woan, G. 2007, *Phys. Rev. D*, 76, 043003
- Courteau, S., et al. 2014, *Rev. Mod. Phys.*, 86, 47
- Daddi, E., et al. 2007, *ApJ*, 670, 156
- Davé, R. 2008, *MNRAS*, 385, 147
- Del Pozzo, W. 2012, *Phys. Rev. D*, 86, 043011
- Dhurandhar, S. V., & Sathyaprakash, B. S. 1994, *Physical Review D*, 49, 1707
- Dominik, M., Belczynski, K., Fryer, C., Holz, D. E., Berti, E., Bulik, T., Mandel, I., & O’Shaughnessy, R. 2013, *ApJ*, 779, 72
- Eichler, D., Livio, M., Piran, T., & Schramm, D. N. 1989, *Nature*, 340, 126
- Evans, P. A., et al. 2012, *ApJS*, 203, 28
- Fairhurst, S. 2011, *Classical and Quantum Gravity*, 28, 105021
- Fan, X. L., Pipino, A., & Matteucci, F. 2013, *ApJ*, 768, 178
- Fan, X. L., Yin, J., & Matteucci, F. 2010, *A&A*, 521, A73
- Fong, W., & Berger, E. 2013, *ApJ*, 776, 18
- Fong, W., et al. 2013, *ApJ*, 769, 56
- Fox, D. B., et al. 2005, *Nature*, 437, 845
- Fryer, C. L., Belczynski, K., Wiktorowicz, G., Dominik, M., Kalogera, V., & Holz, D. E. 2012, *ApJ*, 749, 91
- Fynbo, J. P. U., Malesani, D., & Jakobsson, P. 2012, *Long Gamma-Ray Burst Host Galaxies and their Environments*, 269–301
- Gao, H., Zhang, B., Wu, X.-F., & Dai, Z.-G. 2013, *Phys. Rev. D*, 88, 043010
- Grover, K., et al. 2013, *ArXiv e-prints*:1310.7454, 1310.7454
- Hanna, C., Mandel, I., & Voudsen, W. 2014, *ApJ*, 784, 8
- Harry, G. M., & LIGO Scientific Collaboration. 2010, *Class. Quant. Grav.*, 27, 084006
- Hayama, K., Desai, S., Mohanty, S. D., Rakhmanov, M., Summerscales, T., & Yoshida, S. 2008, *Classical and Quantum Gravity*, 25, 184016
- Jarrett, T. H., Chester, T., Cutri, R., Schneider, S. E., & Huchra, J. P. 2003, *AJ*, 125, 525
- Kanner, J., Camp, J., Racusin, J., Gehrels, N., & White, D. 2012, *ApJ*, 759, 22

- Kanner, J., Huard, T. L., Márka, S., Murphy, D. C., Piscionere, J., Reed, M., & Shawhan, P. 2008, *Classical and Quantum Gravity*, 25, 184034
- Karachentsev, I. D., Makarov, D. I., & Kaisina, E. I. 2013, *AJ*, 145, 101
- Koyama, Y., et al. 2013, *MNRAS*, 434, 423
- Leibler, C. N., & Berger, E. 2010, *ApJ*, 725, 1202
- Maraston, C. 1998, *MNRAS*, 300, 872
- Metzger, B. D., & Berger, E. 2012, *ApJ*, 746, 48
- Metzger, B. D., et al. 2010, *MNRAS*, 406, 2650
- Nissanke, S., Kasliwal, M., & Georgieva, A. 2013, *ApJ*, 767, 124
- Nuttall, L. K., & Sutton, P. J. 2010, *Phys. Rev. D*, 82, 102002
- Nuttall, L. K., White, D. J., Sutton, P. J., Daw, E. J., Dhillon, V. S., Zheng, W., & Akerlof, C. 2013, *ApJS*, 209, 24
- O’Shaughnessy, R., Kalogera, V., & Belczynski, K. 2010, *ApJ*, 716, 615
- Paczynski, B. 1986, *ApJ*, 308, L43
- Schutz, B. F. 1986, *Nature*, 323, 310
- Searle, A. C., Sutton, P. J., Tinto, M., & Woan, G. 2008, *Classical and Quantum Gravity*, 25, 114038
- Sidery, T., et al. 2013, *ArXiv e-prints*:1312.6013, 1312.6013
- Singer, L. P., et al. 2014, *ArXiv e-prints*:1404.5623, 1404.5623
- Sylvestre, J. 2003, *ApJ*, 591, 1152
- The Virgo Collaboration. 2009, technical Report VIR-0027A-09
- Veitch, J., & Vecchio, A. 2010, *Phys. Rev. D*, 81, 062003
- Wen, L., & Chen, Y. 2010, *Phys. Rev. D*, 81, 082001
- Wen, L., Fan, X., & Chen, Y. 2008, *Journal of Physics Conference Series*, 122, 012038
- Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, *ApJ*, 754, L29
- White, D. J., Daw, E. J., & Dhillon, V. S. 2011, *Classical and Quantum Gravity*, 28, 085016